数理解析研究所講究録
第 842 巻 1993 年 36-48

Two-Stage Point Estimation with a Shrinkage Stopping Rule

University of Tokyo, Tatsuya Kubokawa (久保川達也)

Consider the problem of estimating a mean vector in a p-variate normal distribution under two-stage sequential sampling schemes. The paper proposes a stopping rule motivated by the James-Stein shrinkage estimator, and shows that the stopping rule and the corresponding shrinkage estimator asymptotically dominate the usual two-stage procedures under a sequence of local alternatives for $p \geq 3$. Also the results of Monte Carlo simulation for average sample sizes and risks of estimators are given and it is revealed that the improvements of the proposed shrinkage procedures are great when a noncentrality parameter is small.

AMS 1991 subject classifications. Primary 62L12; secondary 62F10.

Key words and phrases. Two-stage procedure; shrinkage estimation; domination of sample size; local alternatives; asymptotic risk.

1. Introduction

Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be a sequence of mutually independent random vectors, $\mathbf{X}_i$ having p-variate normal distribution $\mathrm{N}_p(\theta, \sigma^2 \mathbf{I}_p)$ where θ is an unknown vector and σ^2 is an unknown scalar. Consider the problem of finding estimator δ of the mean vector θ such that for a pre-assigned constant $\varepsilon > 0$,

$$\mathrm{R}(\omega, \delta) \equiv \mathrm{E}_\omega[\|\delta - \theta\|^2] \leq \varepsilon \tag{1.1}$$

uniformly for unknown parameters $\omega = (\theta, \sigma^2)$, where $\|\cdot\|$ denotes the Eucledian norm. This subject requires the boundedness of the risk function and it may be called *a bounded risk problem.*

When sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ for fixed size n is taken, the MLE of θ is the sample mean $\overline{\mathbf{X}}_n = n^{-1} \sum_{i=1}^n \mathbf{X}_i$ with risk $\mathrm{R}(\omega, \overline{\mathbf{X}}_n) = p\sigma^2/n$. If σ^2 is known, the risk is equal to ε for

$$n = n_0 \equiv p\sigma^2/\varepsilon. \tag{1.2}$$

For simplicity, we shall, henceforth assume n_0 to be a positive integer. For unknown σ^2, however, there does not exist any fixed sample size that satisfies (1.1) for all ω. Motivated from (1.2), the following two-stage sampling rule is then proposed (Rao(1973)):

(i) Start with an initial sample $\mathbf{X}_1, \ldots, \mathbf{X}_m$ of size $m(> 1 + 2/p)$.

(ii) Define the stopping number by

$$N_0 = \max\{m,\ [p\sigma_m^2/\varepsilon]\} \tag{1.3}$$

where $[u]$ is the smallest integer $\geq u$ and

$$\sigma_m^2 \equiv \sum_{i=1}^{m} \|\mathbf{X}_i - \overline{\mathbf{X}}_m\|^2/\{p(m-1)-2\}. \tag{1.4}$$

(iii) Take another sample $\mathbf{X}_{m+1}, \ldots, \mathbf{X}_{N_0}$.
We estimate θ by $\overline{\mathbf{X}}_{N_0}$, which satisfies (1.1) because

$$\mathrm{R}(\omega, \overline{\mathbf{X}}_{N_0}) = p\sigma^2 \mathrm{E}[N_0^{-1}] \leq \varepsilon \mathrm{E}[\sigma^2/\sigma_m^2] = \varepsilon.$$

Similarly to Mukhopadhyay(1980) and Ghosh and Mukhopadhyay(1981), it can be easily checked that if $m = O(\varepsilon^{-d})$ for $0 < d < 1$, then the following asymptotic properties hold:

$$\lim_{\varepsilon \to 0} \mathrm{E}[N_0]/n_0 = 1 \qquad \text{(asymptotic efficiency)}, \tag{1.5}$$

$$\lim_{\varepsilon \to 0} \mathrm{R}(\omega, \overline{\mathbf{X}}_{N_0})/\varepsilon = 1 \qquad \text{(asymptotic consistency)}. \tag{1.6}$$

For fixed sample sizes, Stein(1956) and James and Stein(1961) established the inadmissibility of the sample mean under the quadratic loss for $p \geq 3$. As an analogous result, Ghosh and Sen(1983) developed James-Stein type estimators dominating the two-stage sequential estimator $\overline{\mathbf{X}}_{N_0}$ for $p \geq 3$. Further, Takada(1984), Ghosh, Nickerson and Sen(1987) and Nickerson(1987) have shown the similar risk dominance results under purely sequential sampling schemes where stopping rules were not made any shrimkage methodology. These assert the improvements only of estimators by shrinkage estimation. It seems, however, that the discussions about the improvements of stopping rules are important in the sequential analysis.

The purpose of this note is to provide a shrinkage stopping rule and the corresponding shrinkage estimator being superior to the usual ones N_0 and $\overline{\mathbf{X}}_{N_0}$. For this, we first consider the James-Stein estimator

$$\delta_n^{JS} = \overline{\mathbf{X}}_n - \frac{(p-2)\sigma^2}{n\|\overline{\mathbf{X}}_n\|^2}\overline{\mathbf{X}}_n$$

for fixed sample size n and known σ^2, which takes the risk

$$\mathrm{R}(\omega, \delta_n^{JS}) = p\frac{\sigma^2}{n} - \mathrm{E}\Big[\frac{(p-2)^2}{\|\overline{\mathbf{X}}_n\|^2}\Big]\Big(\frac{\sigma^2}{n}\Big)^2$$

as shown by James and Stein(1961). Here we like to incorporate shrinkage factors in the stopping rule as well with a view to establish the dominance along with a smaller average sample size. Since

$$\mathrm{E}[\frac{\sigma^2}{n\|\overline{\mathbf{X}}_n\|^2}] \geq \frac{\sigma^2}{\mathrm{E}[n\|\overline{\mathbf{X}}_n\|^2]} = \frac{1}{p+n\|\theta\|^2/\sigma^2},$$

it is seen that

$$\mathrm{R}(\omega, \delta_n^{JS}) \leq p\frac{\sigma^2}{n} - \frac{(p-2)^2}{p+n\lambda}\frac{\sigma^2}{n} \tag{1.7}$$

where $\lambda = \|\theta\|^2/\sigma^2$. Hence $\mathrm{R}(\omega, \delta_n^{JS}) \leq \varepsilon$ if

$$p\frac{\sigma^2}{n} - \frac{(p-2)^2}{p+n\lambda}\frac{\sigma^2}{n} \leq \varepsilon,$$

or

$$\frac{\varepsilon}{\sigma^2}\lambda n^2 + (\frac{\varepsilon}{\sigma^2} - \lambda)pn - 4(p-1) \geq 0. \tag{1.8}$$

The minimum n satisfying (1.8) is given by

$$n_1 = \frac{(\lambda - \varepsilon/\sigma^2)p + \sqrt{(\lambda - \varepsilon/\sigma^2)^2 p^2 + 16(p-1)\lambda\varepsilon/\sigma^2}}{2\lambda\varepsilon/\sigma^2}, \tag{1.9}$$

and it can be easily checked that

$$\frac{4(p-1)}{p}\frac{\sigma^2}{\varepsilon} \leq n_1 \leq p\frac{\sigma^2}{\varepsilon} = n_0. \tag{1.10}$$

Since σ^2 and λ are unknown, they must be estimated. Kubokawa, Robert and Saleh(1993) proposed, for example, the estimator λ_m of λ as

$$\lambda_m = \max\left\{\frac{\|\overline{\mathbf{X}}_m\|^2}{\sigma_m^2} - \frac{p}{m}, \frac{2\|\overline{\mathbf{X}}_m\|^2}{(p+2)\sigma_m^2}\right\} \tag{1.11}$$

for σ_m^2 defined by (1.4).

Using these estimators σ_m^2 and λ_m, we can define a stopping rule N motivated by the James-Stein shrinkage rule, by

$$N = \max\{m, [N_1]\}, \tag{1.12}$$

where

$$N_1 = \frac{(\lambda_m - \varepsilon/\sigma_m^2)p + \sqrt{(\lambda_m - \varepsilon/\sigma_m^2)^2 p^2 + 16(p-1)\lambda_m\varepsilon/\sigma_m^2}}{2\lambda_m\varepsilon/\sigma_m^2}. \tag{1.13}$$

Based on the sample $\mathbf{X}_1, \ldots, \mathbf{X}_N$, θ is estimated by

$$\delta_N = \overline{\mathbf{X}}_N - \frac{a\sigma_N^2}{N\|\overline{\mathbf{X}}_N\|^2}\overline{\mathbf{X}}_N \tag{1.14}$$

where a is a positive constant suitably chosen. Similar to (1.10), we observe

$$\frac{4(p-1)}{p}\frac{\sigma_m^2}{\varepsilon} \leq N_1 \leq p\frac{\sigma_m^2}{\varepsilon}, \tag{1.15}$$

which implies that

$$N \leq N_0 \quad \text{a.s.},$$

so that

$$\mathrm{E}[N] \leq \mathrm{E}[N_0] \tag{1.16}$$

for all ω. This means that N has an exactly smaller average sample size than N_0. On the other hand, it may be difficult to evaluate the risk function of δ_N exactly. Thereby, we discuss the asymptotic properties of N and δ_N.

Section 2 presents the asymptotic efficiency of N and the consistency of δ_N, that is, $\lim_{\varepsilon\to 0}\mathrm{E}[N]/n_0 = 1$ and $\lim_{\varepsilon\to 0}\mathrm{R}(\omega, \delta_N)/\varepsilon = 1$ for uniformly $\theta \neq 0$ and $p \geq 3$. From (1.5) and (1.6), they mean that N and δ_N are asymptotically equivalent to N_0 and $\overline{\mathbf{X}}_{N_0}$.

In Section 3, to compare them for θ close to zero, we consider a sequence of local alternatives $\theta = \theta_{n_0} = n_0^{-1/2}\theta_0$ for fixed θ_0. Under the local alternatives, we get that $\lim_{\varepsilon\to 0}\mathrm{E}[N]/n_0 = 1$ and for $p \geq 3$,

$$\lim_{\varepsilon\to 0}\mathrm{R}(\omega, \delta_N)/\varepsilon = 1 - \frac{a}{p}\mathrm{E}\Big[\frac{2(p-2)-a}{\chi_p^2(\lambda_0)}\Big], \tag{1.17}$$

where $\chi_p^2(\lambda_0)$ denotes a noncentral chi square variate with p degrees of freedom and noncentrality parameter $\lambda_0 = \|\theta_0\|^2/\sigma^2$. From (1.6) and (1.17), we can see that, if $0 < a < 2(p-2)$ then

$$\lim_{\varepsilon\to 0}\{\mathrm{R}(\omega, \delta_N) - \mathrm{R}(\omega, \overline{\mathbf{X}}_{N_0})\}/\varepsilon < 0,$$

that is, δ_N dominates $\overline{\mathbf{X}}_{N_0}$ asymptotically for $p \geq 3$. Also (1.17) means that δ_N dominates $\overline{\mathbf{X}}_N$ asymptotically for the same shrinkage stopping rule N. In this way, we obtain the shrinkage procedures N and δ_N such that N is exactly smaller than or equal to N_0 and δ_N is asymptotically better than $\overline{\mathbf{X}}_{N_0}$ and $\overline{\mathbf{X}}_N$ for $p \geq 3$.

The results of Monte Carlo simulation for the average sample sizes $\mathrm{E}[N_0]$, $\mathrm{E}[N]$ and the risks $\mathrm{R}(\omega, \overline{\mathbf{X}}_{N_0})$, $\mathrm{R}(\omega, \delta_N)$ are given in Section 4. This is done in the cases of $m = 5$; $p = 4, 8$; $\varepsilon = 1.0, 0.5, 0.3$, and it is revealed that the improvements of N and δ_N are great when noncentrality parameter $\|\theta\|^2/\sigma^2$ is small.

2. Asymptotic efficiency and consistency for fixed alternatives

We shall investigate the asymptotic properties of N and δ_N given by (1.12) and (1.14) for any fixed $\theta \neq 0$.

Theorem 2.1. Assume that $m = O(\varepsilon^{-d})$ for $0 < d < 1$. Then for uniformly $\theta \neq 0$,

(i) $\lim_{\varepsilon \to 0} \mathrm{E}[N]/n_0 = 1$,

(ii) $\lim_{\varepsilon \to 0} \mathrm{R}(\omega, \delta_N)/\varepsilon = 1$ for $p \geq 3$.

For the proof, the following lemmas are essential.

Lemma 2.1. Assume that $m = O(\varepsilon^{-d})$ for $0 < d < 1$. Then,

(i) $\lim_{\varepsilon \to 0} N/n_0 = 1$ a.s.,

(ii) N/n_0 is uniformly integrable,

(iii) $(n_0/N)^\alpha$ is uniformly integrable for fixed $\alpha > 0$.

Proof. From the definition (1.12),

$$\frac{\lambda_m \sigma_m^2 - \varepsilon + \sqrt{(\lambda_m \sigma_m^2 - \varepsilon)^2 + 16(p-1)\varepsilon \lambda_m \sigma_m^2/p^2}}{2\sigma^2 \lambda_m} \leq \frac{N}{n_0} <$$
$$\frac{\lambda_m \sigma_m^2 - \varepsilon + \sqrt{(\lambda_m \sigma_m^2 - \varepsilon)^2 + 16(p-1)\varepsilon \lambda_m \sigma_m^2/p^2}}{2\sigma^2 \lambda_m} + \frac{m+1}{n_0}. \tag{2.1}$$

From the condition of the lemma it is easy to see that $m \to \infty$ and $(m+1)/n_0 \to 0$ as $\varepsilon \to 0$. Also, $\sigma_m^2 \to \sigma^2$ a.s. and $\lambda_m \to \lambda$ a.s. as $m \to \infty$. Then part (i) follows from (2.1). For any d and $\delta > 0$, we observe that

$$\mathrm{E}\Big[\frac{N}{n_0} \mathrm{I}_{[N/n_0 > d]}\Big] \leq d^{-\delta} \mathrm{E}\Big[\Big(\frac{N}{n_0}\Big)^{1+\delta}\Big], \tag{2.2}$$

where $\mathrm{I}_{[\cdot]}$ denotes the indicator function. To prove (ii), it suffices to show

$$\sup\nolimits_{0<\varepsilon<\varepsilon_0} \{\mathrm{E}[(N/n_0)^{1+\delta}]\} < \infty$$

for fixed $\varepsilon_0 > 0$. From (1.15) and (2.1),

$$\mathrm{E}\Big[\Big(\frac{N}{n_0}\Big)^{1+\delta}\Big] \leq K_0 \Big\{\mathrm{E}\Big[\sup_{m \geq 2} (\sigma_m^2/\sigma^2)^{1+\delta}\Big] + \Big(\frac{m+1}{n_0}\Big)^{1+\delta}\Big\} \tag{2.3}$$

for some constant K_0 independent of ε. By Doob's maximal inequality for the reversed martingale sequence, $\mathrm{E}[\sup_{m \geq 2} (\sigma_m^2/\sigma^2)^{1+\delta}] < \infty$. Also $(m+1)/n_0 \to 0$ as $\varepsilon \to 0$, so that $\sup_{0<\varepsilon<\varepsilon_0} \{\mathrm{E}[(N/n_0)^{1+\delta}]\} < \infty$, which proves (ii). For part (iii), similar to (2.2), we show that

$$\sup_{0<\varepsilon<\varepsilon_0} \{\mathrm{E}[(n_0/N)^{\alpha(1+\delta)}]\} < \infty. \tag{2.4}$$

From (1.15) and (2.1), for some constant $K_1 > 0$,

$$\begin{aligned}
&\mathrm{E}[(\frac{n_0}{N})^{\alpha(1+\delta)}] \\
&\leq K_1 \mathrm{E}[(\frac{\sigma^2}{\sigma_m^2})^{\alpha(1+\delta)}] \\
&= K_1\{p(m-1)-2\}^{\alpha(1+\delta)} \frac{\Gamma(p(m-1)/2-\alpha(1+\delta))}{\Gamma(p(m-1)/2)} 2^{-\alpha(1+\delta)},
\end{aligned} \tag{2.5}$$

which is bounded by a constant independent of ε and the proof is complete.

Lemma 2.2. Assume that $m = O(\varepsilon^{-d})$ for $0 < d < 1$. Then, for $p \geq 3$,
(i) $n_0\|\overline{\mathbf{X}}_N - \theta\|^2$ is uniformly integrable,
(ii) $n_0\sigma_N^4/(N^2\|\overline{\mathbf{X}}_N\|^2)$ is uniformly integrable,
(iii) $n_0\sigma_N^2\overline{\mathbf{X}}_N'(\overline{\mathbf{X}}_N - \theta)/(N\|\overline{\mathbf{X}}_N\|^2)$ is uniformly integrable.

Proof. Similar to (2.2), it suffices to show that for $\delta > 0$,

$$\sup_{0<\varepsilon<\varepsilon_0} \left\{\mathrm{E}[(n_0\|\overline{\mathbf{X}}_N - \theta\|^2)^{1+\delta}]\right\} < \infty. \tag{2.6}$$

By Hölder's inequality, for small $\delta' > 0$,

$$\begin{aligned}
\mathrm{E}[(n_0\|\overline{\mathbf{X}}_N - \theta\|^2)^{1+\delta}] \leq &\left\{\mathrm{E}[(\frac{n_0}{N})^{(1+\delta)(1+\delta')}]\right\}^{1/(1+\delta')} \\
&\times \left\{\mathrm{E}[(N\|\overline{\mathbf{X}}_N - \theta\|^2)^{(1+\delta)(1+\delta')/\delta'}]\right\}^{\delta'/(1+\delta')}
\end{aligned}$$

From (iii) of Lemma 2.1, $\sup_{0<\varepsilon<\varepsilon_0}\mathrm{E}[(n_0/N)^{(1+\delta)(1+\delta')}] < \infty$, and

$$\begin{aligned}
\mathrm{E}[(N\|\overline{\mathbf{X}}_N - \theta\|^2)^{(1+\delta)(1+\delta')/\delta'}] &\leq \mathrm{E}[\sup_{n\geq 2}(n\|\overline{\mathbf{X}}_n - \theta\|^2)^{(1+\delta)(1+\delta')/\delta'}] \\
&= \mathrm{E}[(\sigma^2\|\mathbf{Z}^*\|^2)^{(1+\delta)(1+\delta')/\delta'}],
\end{aligned}$$

which is also bounded, where

$$\mathbf{Z}^* = \sqrt{n}(\overline{\mathbf{X}}_n - \theta)/\sigma \tag{2.7}$$

is $\mathrm{N}_p(\mathbf{0}, \mathbf{I}_p)$ independent of n. Hence we get (2.6) and prove (i). For part (ii), we shall prove

$$\sup_{0<\varepsilon<\varepsilon_0}\left\{\mathrm{E}\Big[\Big(\frac{n_0\sigma_N^4}{N^2\|\overline{\mathbf{X}}_N\|^2}\Big)^{1+\delta}\Big]\right\} < \infty. \tag{2.8}$$

By Hölder's inequality,

$$\begin{aligned}
\mathrm{E}\Big[\Big(\frac{n_0\sigma_N^4}{N^2\|\overline{\mathbf{X}}_N\|^2}\Big)^{1+\delta}\Big] \leq &\left\{\mathrm{E}\Big[\Big(\frac{n_0}{N^2\|\overline{\mathbf{X}}_N\|^2}\Big)^{(1+\delta)(1+\delta')}\Big]\right\}^{1/(1+\delta')} \\
&\times \left\{\mathrm{E}[\sigma_N^{2(1+\delta)(1+\delta')/\delta'}]\right\}^{\delta'/(1+\delta')}.
\end{aligned} \tag{2.9}$$

For $\gamma = (1+\delta)(1+\delta')$ with small $\delta' > 0$, clearly,

$$\mathrm{E}[\sigma_N^{2\gamma/\delta'}] \le \mathrm{E}[\sup_{n\ge 2}(\sigma_n^{2\gamma/\delta'})] < \infty.$$

For small $\delta'' > 0$,

$$\begin{aligned}
&\mathrm{E}[(\frac{n_0}{N^2\|\overline{\mathbf{X}}_N\|^2})^\gamma] \\
&\le \left\{\mathrm{E}[(\frac{n_0}{N})^{\gamma(1+\delta'')/\delta''}]\right\}^{\delta''/(1+\delta'')}\left\{\mathrm{E}[(\frac{1}{N\|\overline{\mathbf{X}}_N\|^2})^{\gamma(1+\delta'')}]\right\}^{1/(1+\delta'')}
\end{aligned} \tag{2.10}$$

From the proof of (iii) of Lemma 2.1, we see that

$$\sup_{0<\varepsilon<\varepsilon_0}\mathrm{E}[(n_0/N)^{\gamma(1+\delta'')/\delta''}] < \infty.$$

By use of $\mathbf{Z}^*$ in (2.7), $n\|\overline{\mathbf{X}}_n\|^2$ is represented as

$$n\|\overline{\mathbf{X}}_n\|^2 = \sigma^2\|\mathbf{Z}^* + \sqrt{n}\theta/\sigma\|^2 = \sigma^2\chi^2_{p+2J_n} = \sigma^2\chi^2_p + \sigma^2\chi^2_{2J_n},$$

where J_n follows a Poisson law with $\mathrm{E}[J_n] = n\|\theta\|^2/(2\sigma^2)$. Since χ^2_p is independent of n,

$$\begin{aligned}
\mathrm{E}[(N\|\overline{\mathbf{X}}_N\|^2)^{-\gamma(1+\delta'')}] &\le \mathrm{E}[\sup_{n\ge 2}(n\|\overline{\mathbf{X}}_n\|^2)^{-\gamma(1+\delta'')}] \\
&= \mathrm{E}[\sup_{n\ge 2}(\sigma^2\chi^2_p + \sigma^2\chi^2_{2J_n})^{-\gamma(1+\delta'')}] \\
&\le \mathrm{E}[(\sigma^2\chi^2_p)^{-\gamma(1+\delta'')}],
\end{aligned} \tag{2.11}$$

which is bounded for $p \ge 3$ for sufficiently small $\delta'' > 0$. Combining (2.9), (2.10) and (2.11) yields part (ii). Since $\overline{\mathbf{X}}'_N(\overline{\mathbf{X}}_N - \theta) \le \|\overline{\mathbf{X}}_N\|\|\overline{\mathbf{X}}_N - \theta\|$, we have that

$$\begin{aligned}
&\mathrm{E}[\left|\frac{n_0\sigma_N^2}{N\|\overline{\mathbf{X}}_N\|^2}\overline{\mathbf{X}}'_N(\overline{\mathbf{X}}_N - \theta)\right|^{1+\delta}] \\
&\le \mathrm{E}[(\frac{\sqrt{n_0}\sigma_N^2}{N\|\overline{\mathbf{X}}_N\|})^{1+\delta}(\sqrt{n_0}\|\overline{\mathbf{X}}_N - \theta\|)^{1+\delta}] \\
&\le \left\{\mathrm{E}[(\frac{n_0\sigma_N^4}{N^2\|\overline{\mathbf{X}}_N\|^2})^{1+\delta}]\right\}^{1/2}\left\{\mathrm{E}[(n_0\|\overline{\mathbf{X}}_N - \theta\|^2)^{1+\delta}]\right\}^{1/2},
\end{aligned}$$

so that part (iii) follows from (2.6) and (2.8).

Proof of Theorem 2.1. From (i) and (ii) of Lemma 2.1, we get $\lim_{\varepsilon\to 0}\mathrm{E}[N]\,/n_0 = 1$, yielding part (i). For the proof of part (ii), we have that

$$\begin{aligned}
\mathrm{R}(\omega, \delta_N) &= \mathrm{R}(\omega, \overline{\mathbf{X}}_N) \\
&\quad + \mathrm{E}[\frac{a^2\sigma_N^4}{N^2\|\overline{\mathbf{X}}_N\|^2} - \frac{2a\sigma_N^2}{N\|\overline{\mathbf{X}}_N\|^2}\overline{\mathbf{X}}'_N(\overline{\mathbf{X}}_N - \theta)],
\end{aligned}$$

or

$$
\begin{aligned}
\mathrm{R}(\omega, \delta_N)/\varepsilon = \frac{1}{p\sigma^2}\mathrm{E}[n_0\|\overline{\mathbf{X}}_N - \theta\|^2 \\
+ \frac{a^2 n_0 \sigma_N^4}{N^2\|\overline{\mathbf{X}}_N\|^2} - \frac{2an_0\sigma_N^2}{N\|\overline{\mathbf{X}}_N\|^2}\overline{\mathbf{X}}'_N(\overline{\mathbf{X}}_N - \theta)].
\end{aligned} \tag{2.12}
$$

We first consider the case of $N > m$. Let

$$
\begin{aligned}
\mathbf{Z} &= \sqrt{N-m}\left(\frac{1}{N-m}\sum_{i=m+1}^{N}\mathbf{X}_i - \theta\right)/\sigma, \qquad (2.13)\\
\mathbf{U}_m &= \sqrt{m}(\overline{\mathbf{X}}_m - \theta)/\sigma,\\
\mathbf{Y}_\varepsilon &= \left(\frac{m}{N}\right)^{1/2}\mathbf{U}_m + \left(\frac{N-m}{N}\right)^{1/2}\mathbf{Z}.
\end{aligned}
$$

Given N, $\mathbf{Z}$ is conditionally $\mathrm{N}_p(\mathbf{0}, \mathbf{I}_p)$, so that $\mathbf{Z}$ is independent of N or $(\overline{\mathbf{X}}_m, \sigma_m^2)$. Since $\sqrt{N}\overline{\mathbf{X}}_N = \sigma(\mathbf{Y}_\varepsilon + \sqrt{N}\theta/\sigma)$ and $\sqrt{N}(\overline{\mathbf{X}}_N - \theta) = \sigma\mathbf{Y}_\varepsilon$, we write

$$
n_0\|\overline{\mathbf{X}}_N - \theta\|^2 = \frac{n_0}{N}\sigma^2\|\mathbf{Y}_\varepsilon\|^2, \tag{2.14}
$$

$$
\frac{n_0\sigma_N^4}{N^2\|\overline{\mathbf{X}}_N\|^2} = \frac{n_0}{N}\frac{\sigma^2(\sigma_N^2/\sigma^2)^2}{\|\mathbf{Y}_\varepsilon + \sqrt{N}\theta/\sigma\|^2}, \tag{2.15}
$$

$$
\frac{n_0\sigma_N^2}{N\|\overline{\mathbf{X}}_N\|^2}\overline{\mathbf{X}}'_N(\overline{\mathbf{X}}_N - \theta) = \frac{n_0}{N}\sigma_N^2\frac{(\mathbf{Y}_\varepsilon + \sqrt{N}\theta/\sigma)'\mathbf{Y}_\varepsilon}{\|\mathbf{Y}_\varepsilon + \sqrt{N}\theta/\sigma\|^2}, \tag{2.16}
$$

which are convergent in distribution to $\sigma^2\|\mathbf{Z}\|^2$, 0 and 0, respectively, as $\varepsilon \to 0$. Hence combining (2.12) and Lemmas 2.1 and 2.2 provides that

$$
\lim_{\varepsilon\to 0}\mathrm{E}[\{\|\delta_N - \theta\|^2/\varepsilon\}\mathrm{I}_{[N>m]}] = 1. \tag{2.17}
$$

Finally, we need to verify that

$$
\lim_{\varepsilon\to 0}\mathrm{E}[\{\|\delta_m - \theta\|^2/\varepsilon\}\mathrm{I}_{[N=m]}] = 0. \tag{2.18}
$$

From (2.12), it is sufficient to show that

$$
\lim_{\varepsilon\to 0}\mathrm{E}[n_0\|\overline{\mathbf{X}}_m - \theta\|^2\mathrm{I}_{[N=m]}] = 0, \tag{2.19}
$$

$$
\lim_{\varepsilon\to 0}\mathrm{E}\left[\frac{n_0\sigma_m^4}{m^2\|\overline{\mathbf{X}}_m\|^2}\mathrm{I}_{[N=m]}\right] = 0, \tag{2.20}
$$

$$
\lim_{\varepsilon\to 0}\mathrm{E}\left[\frac{n_0\sigma_m^2}{m\|\overline{\mathbf{X}}_m\|^2}\overline{\mathbf{X}}'_m(\overline{\mathbf{X}}_m - \theta)\mathrm{I}_{[N=m]}\right] = 0. \tag{2.21}
$$

Noting that $\{N = m\} = \{N_1 \leq m\} \subset \{\sigma_m^2 \leq \varepsilon mp/4(p-1)\}$ by (1.15), from the Chebyshev inequality, we observe that

$$\begin{aligned}
\mathrm{E}[\mathrm{I}_{[N=m]}] &\leq \mathrm{P}\Big[(\frac{1}{\sigma_m^2})^2 \geq (\frac{4(p-1)}{\varepsilon mp})^2\Big] \\
&\leq (\frac{\varepsilon mp}{4(p-1)})^2 \mathrm{E}[(\frac{1}{\sigma_m^2})^2] \\
&= O(\varepsilon^{2(1-d)})
\end{aligned}$$

for large m or small ε. Then applying Hölder's inequality gives that for $0 < \delta < 1$,

$$\begin{aligned}
&\mathrm{E}[n_0\|\overline{\mathbf{X}}_m - \theta\|^2 \mathrm{I}_{[N=m]}] \\
&\leq \frac{n_0}{m}\Big\{\mathrm{E}[(m\|\overline{\mathbf{X}}_m - \theta\|^2)^{(1+\delta)/\delta}]\Big\}^{\delta/(1+\delta)}\Big\{\mathrm{E}[\mathrm{I}_{[N=m]}]\Big\}^{1/(1+\delta)} \\
&= O(\varepsilon^{(1-d)(2/(1+\delta)-1)}),
\end{aligned}$$

which proves (2.19). For (2.20),

$$\begin{aligned}
\mathrm{E}[\frac{n_0\sigma_m^4}{m^2\|\overline{\mathbf{X}}_m\|^2}\mathrm{I}_{[N=m]}] &\leq \mathrm{E}[\frac{n_0\sigma_m^4}{m^2\|\overline{\mathbf{X}}_m\|^2}\mathrm{I}_{[\sigma_m^2 \leq \varepsilon mp/4(p-1)]}] \\
&\leq (\frac{p}{4(p-1)})^2\frac{n_0}{m}(\varepsilon m)^2\mathrm{E}[\frac{1}{m\|\overline{\mathbf{X}}_m\|^2}] \\
&\leq (\frac{p}{4(p-1)})^2\frac{n_0}{m}(\varepsilon m)^2\frac{1}{(p-2)\sigma^2} \\
&= O(\varepsilon^{1-d}).
\end{aligned}$$

For (2.21), observe that

$$\begin{aligned}
&\mathrm{E}[\frac{n_0\sigma_m^2}{m\|\overline{\mathbf{X}}_m\|^2}\overline{\mathbf{X}}_m'(\overline{\mathbf{X}}_m - \theta)\mathrm{I}_{[N=m]}] \\
&\leq \mathrm{E}[\frac{\sqrt{n_0}\sigma_m^2}{m\|\overline{\mathbf{X}}_m\|}\sqrt{n_0}\|\overline{\mathbf{X}}_m - \theta\|\mathrm{I}_{[N=m]}] \\
&\leq \Big\{\mathrm{E}[\frac{n_0\sigma_m^4}{m^2\|\overline{\mathbf{X}}_m\|^2}\mathrm{I}_{[N=m]}]\Big\}^{1/2}\Big\{\mathrm{E}[n_0\|\overline{\mathbf{X}}_m - \theta\|^2\mathrm{I}_{[N=m]}]\Big\}^{1/2}.
\end{aligned}$$

Then combining (2.19) and (2.20) gives (2.21). Therefore the proof of Theorem 2.1 is complete.

From the proof of Theorem 2.1, we can see that δ_N and $\overline{\mathbf{X}}_N$ for the shrinkage stopping rule N are asymptotically risk-equivalent for fixed $\theta \neq 0$.

3. Asymptotic domination under local alternatives

To compare asymptotically the usual two-stage procedure and the proposed shrinkage procedure for θ close to zero, we consider a sequence of the local alternatives:

$$\theta = \theta_{n_0} = n_0^{-1/2}\theta_0 \qquad \text{for fixed } \theta_0 \in \mathrm{R}^p, \tag{3.1}$$

where n_0 is defined by (1.2).

Theorem 3.1. Assume that $m = O(\varepsilon^{-d})$ for $0 < d < 1$. Then under the local alternatives (3.1),

(i) $\lim_{\varepsilon\to 0}\mathrm{E}[N]/n_0 = 1$,

(ii) $\lim_{\varepsilon\to 0}\mathrm{R}(\omega, \delta_N)/\varepsilon = 1 - \mathrm{E}[a(2(p-2)-a)/\{p\chi_p^2(\lambda_0)\}]$, for $p \geq 3$, where $\chi_p^2(\lambda_0)$ designates a noncentral chi square random variate with degrees of freedom p and noncentrality parameter $\lambda_0 = \|\theta_0\|^2/(2\sigma^2)$.

Proof. Letting $\mathbf{Y} = \sqrt{m}(\overline{\mathbf{X}}_m - \theta_{n_0})/\sigma$, we can see that $\mathbf{Y}$ has $\mathrm{N}_p(\mathbf{0}, \mathbf{I}_p)$ and that $m\|\overline{\mathbf{X}}_m\|^2 = \sigma^2\|\mathbf{Y} + \sqrt{m}\theta_{n_0}/\sigma\|^2$, which is convergent to $\sigma^2\|\mathbf{Y}\|^2$ as $\varepsilon \to 0$ by (3.1) and the fact that $m/n_0 \to 0$ as $\varepsilon \to 0$. Then from (2.1), we can show that $N/n_0 \to 1$ a.s. as $\varepsilon \to 0$. The uniform integrability of N/n_0 can be verified by the same arguments as in the proof of (i) of Lemma 2.1, and part (i) of Theorem 3.1 is proved. For part (ii), recall that $\mathbf{Z}$ defined by (2.13) has $\mathrm{N}_p(\mathbf{0}, \mathbf{I}_p)$ independent of $(\overline{\mathbf{X}}_m, \sigma_m^2)$. On the set $\{N > m\}$, from (2.14), (2.15) and (2.16), it follows that as $\varepsilon \to 0$,

$$\begin{aligned}
n_0\|\overline{\mathbf{X}}_N - \theta\|^2 &\to \sigma^2\|\mathbf{Z}\|^2 \quad \text{a.s.} \\
\frac{n_0\sigma_N^4}{N^2\|\overline{\mathbf{X}}_N\|^2} &\to \frac{\sigma^2}{\|\mathbf{Z} + \theta_0/\sigma\|^2} \quad \text{a.s.} \\
\frac{n_0\sigma_N^2}{N\|\overline{\mathbf{X}}_N\|^2}\overline{\mathbf{X}}_N'(\overline{\mathbf{X}}_N - \theta) &\to \frac{\sigma^2}{\|\mathbf{Z} + \theta_0/\sigma\|^2}(\mathbf{Z} + \theta_0/\sigma)'\mathbf{Z} \quad \text{a.s..}
\end{aligned}$$

Here it will be noted that the uniform integrabilities given by Lemmas 2.1 and 2.2 can be verified by the same arguments for the local alternatives (3.1). Also (2.18) can be shown similarly. Hence from (2.12),

$$\begin{aligned}
\mathrm{R}(\omega, \delta_N)/\varepsilon \to \frac{1}{p}\mathrm{E}\Big[&\|\mathbf{Z}\|^2 + \frac{a^2}{\|\mathbf{Z} + \theta_0/\sigma\|^2} \\
&- \frac{2a}{\|\mathbf{Z} + \theta_0/\sigma\|^2}(\mathbf{Z} + \theta_0/\sigma)'\mathbf{Z}\Big]
\end{aligned} \tag{3.2}$$

as $\varepsilon \to 0$. Applying the Stein identity given by Stein(1981) to the r.h.s. of (3.2) gives the expression of Theorem 3.1 (ii), and the proof is complete.

From (1.6), (1.16) and Theorem 3.1, we can see

Corollary 3.1. Assume that $m = O(\varepsilon^{-d})$ for $0 < d < 1$. If $p \geq 3$ and $0 < a < 2(p-2)$, then under the local alternatives (3.1), δ_N dominates $\overline{\mathbf{X}}_{N_0}$ asymptotically while N dominates N_0 exactly. Also δ_N asymptotically dominates $\overline{\mathbf{X}}_N$ for the same shrinkage stopping rule N.

4. Simulation results

In this section we present the results of Monte Carlo simulation for the average sample sizes $\mathrm{E}[N_0]$, $\mathrm{E}[N]$ and the risks $\mathrm{R}(\omega, \overline{\mathbf{X}}_{N_0})$, $\mathrm{R}(\omega, \delta_N)$. This is done in the cases of $m = 5$; $p = 4, 8$; $\varepsilon = 1.0, 0.5, 0.3$; $\|\theta\|^2 = 0.0, 0.5, 1.0, 1.5$; $\sigma^2 = 0.5, 1.0, 2.0$. Tables 1, 2 and Tables 3, 4 report the average values, respectively, of the stopping numbers N_0, N and of the losses of the estimators $\overline{\mathbf{X}}_{N_0}$, δ_N based on 20000 replications. From the tables, we see that N and δ_N are relatively superior to N_0 and $\overline{\mathbf{X}}_{N_0}$ and that their improvements are great when noncentrality parameter $\|\theta\|^2/\sigma^2$ is small. When $m = 5$, $p = 8$, $\varepsilon = 0.3$, $\|\theta\|^2 = 0.0$ and $\sigma^2 = 2.0$, for example, the gain in sampling $\mathrm{E}[N_0] - \mathrm{E}[N]$ is 12.8 and the relative risk improvement $100 \times \{\mathrm{R}(\omega, \overline{\mathbf{X}}_{N_0}) - \mathrm{R}(\omega, \delta_N)\}/\mathrm{R}(\omega, \overline{\mathbf{X}}_{N_0})$ is 69%.

Acknowledgement. The authors wish to thank Professor P.K. Sen for his valuable comments. Also they are grateful to Mr. T. Yoshiba for the help in numerical computation.

REFERENCES

Ghosh, M. and Mukhopadhyay, N.(1981). Consistency and asymptotic efficiency of two-stage and sequential estimation procedures. *Sankhya, A*, **43**, 220-227.

Ghosh, M., Nickerson, D.M. and Sen, P.K.(1987). Sequential shrinkage estimation. *Ann. Statist.*, **15**, 817-829.

Ghosh, M. and Sen, P.K.(1983). On two-stage James-Stein estimators. *Sequential Anal.*, **2**, 359-367.

James, W. and Stein, C.(1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.*, **1**, 361-379.

Kubokawa, T., Robert, C.P. and Saleh, A.K.Md.E.(1993). Estimation of the noncentrality parameters. *Canadian J. Statist.*, **21**, 45-57.

Mukhopadhyay, N.(1980). A consistent and asymptotically efficient two-stage procedure to construct fixed width confidence intervals for the mean. *Metrika*, **27**, 281-284.

Nickerson, D.M.(1987). Sequential shrinkage estimation of linear regression parameters. *Sequential Anal.*, **6**, 93-117.

Rao, C.R.(1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.

Stein, C.(1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Prob.*, **1**, 197-206.

Stein, C.(1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.*, **9**, 1135-1151.

Takada, Y.(1984). Inadmissibility of a sequential estimation rule of the mean of a multivariate normal distribution. *Sequential Anal.*, **3**, 267-271

Table 1. Average sample sizes $E[N_0]$ and $E[N]$ for $m = 5$ and $p = 4$.

ε		1.0			0.5			0.3		
	$\|\|\theta\|\|^2$ \ σ^2	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
N	0.0	5.0	5.2	8.0	5.2	8.0	16.2	6.9	13.4	27.5
N	0.5	5.0	5.2	8.2	5.2	8.3	16.5	7.2	13.9	28.1
N	1.0	5.0	5.2	8.3	5.3	8.6	16.8	7.5	14.4	28.5
N	1.5	5.0	5.3	8.5	5.4	8.8	17.0	7.7	14.6	28.8
N_0	-	5.0	5.6	9.6	5.6	9.6	18.7	8.1	15.7	31.0

Table 2. Average sample sizes $E[N_0]$ and $E[N]$ for $m = 5$ and $p = 8$.

ε		1.0			0.5			0.3		
	$\|\|\theta\|\|^2$ \ σ^2	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
N	0.0	5.0	5.5	10.8	5.4	10.8	24.5	8.8	19.7	44.6
N	0.5	5.0	5.7	11.3	6.0	11.8	25.5	10.3	21.4	46.2
N	1.0	5.0	6.0	11.8	6.7	12.7	26.5	11.7	22.8	47.6
N	1.5	5.0	6.4	12.3	7.2	13.5	27.3	12.5	24.1	48.9
N_0	-	5.3	9.0	17.5	9.0	17.5	34.5	14.7	28.9	57.4

Table 3. Risks of estimators $\overline{\mathbf{X}}_{N_0}$ and δ_N for $m = 5$ and $p = 4$.

ε		1.0			0.5			0.3		
	$\|\|\theta\|\|^2$ \ σ^2	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
δ_N	0.0	.24	.43	.53	.21	.27	.27	.16	.16	.15
δ_N	0.5	.34	.57	.71	.32	.40	.40	.24	.26	.26
δ_N	1.0	.36	.64	.81	.34	.43	.44	.26	.28	.28
δ_N	1.5	.37	.68	.84	.35	.45	.46	.26	.28	.29
$\overline{\mathbf{X}}_{N_0}$	-	.40	.73	.91	.36	.46	.48	.26	.28	.29

Table 4. Risks of estimators $\overline{\mathbf{X}}_{N_0}$ and δ_N for $m = 5$ and $p = 8$.

ε		1.0			0.5			0.3		
	$\|\|\theta\|\|^2$ \ σ^2	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
δ_N	0.0	.24	.40	.36	.20	.18	.16	.11	.10	.09
δ_N	0.5	.47	.66	.64	.40	.39	.38	.28	.27	.25
δ_N	1.0	.57	.81	.80	.47	.48	.46	.29	.30	.29
δ_N	1.5	.62	.89	.91	.48	.50	.49	.30	.31	.30
$\overline{\mathbf{X}}_{N_0}$	-	.76	.93	.97	.46	.48	.49	.28	.29	.29